RDF for dummies

Семантика

С этой статьи я начинаю совй цикл постов «для новичков» где максимально популярно растолкую понятия веб 3.0. В последствии все статьи перекочуют в вики и будут «изданы» мною в виде PDF книги.

Начнем со средств, и сегодня у нас основа основ — RDF.

Resource Description Framework — это разработанная консорциумом W3C модель для описания ресурсов, в особенности — метаданных о ресурсах.

RDF это язык описания знаний. Это не совсем XML. то есть совсем не XML, просто синтаксис похож. RDF содержит тройки данных «объект — предикат — субъект». ну пример «Столб имеетВысоту 15м». Вот простейший пример; RDF.:

```
@prefix : <http: www.example.org> .
:john a :Person .
:john :hasMother :helga .
:john :hasFather :henrich .
:richard :hasSister :jane .
```

Назовем это Документ #1, он потом нам пригодиться.

Пока понятно? Нет?

```
@prefix - подгружаем пространство имен (тут хранятся условные всякие
oписания, которые понятны компьютеру, потом расскажу по-подробнее)
:john a :Person — некто Джон человек (безусловно)
:john :hasMother : helga — У Джона есть мама и её зовут Хельга.
:john :hasFather : henrich — а отца Джона зовут Генрих
:richard :has Sister : jane — у отца есть сестра — Джейн
```

Такая вот себе семейка. Кстати, это короткая форма записи RDF, зовется она N3 (Нотация три). Как я понимаю, создана под влиянием ЛИСПа и после того, как поняли, что XML слишком большой. Она хороша в некоторых случаях и очень наглядна, будем использовать её в качестве примера. Но есть и XML-like версия, она и используется в вебе — RSS 0.9, (если я не ошибаюсь) тому пример. Вот XML-like пример про того же Джона:

```
<rdf:rdf xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:ns="http://www.example.org/#john">
<ns:hasmother rdf:resource="http://www.example.org/#helga" />
<ns:hasfather>
```

```
<rdf:description rdf:about="http://www.example.org/#henrich">
  <ns:hassister rdf:resource="http://www.example.org/#jane">
    </ns:hassister>
    </rdf:description>
    </ns:hasfather>
  </ns:person>
  </rdf:rdf>
```

Малопонятно и плохочитаемо. Согласен, но для компьютера удобнее, так как любой язык программирования может с легкостью разобрать XML-файл и получить данные.

Самое сильное отличие RDF от XML заключается в том, что RDF предназначен для распределенных данных. К примеру мой SIOC. Я перепечатывал посты Джона Бреслина, значит мой SIOC связан с его SIOCом и компьютер может легко, по ссылкам (ссылки главная опора Семантической Сети, помните?), собрать единое обсуждение. Роботы могут собрать разные RDF, написанные разными людьми и узнать вещи, которых не было явно ни в одном из документов, чувствуете приближение Матрицы?

Пример:

Из Документа #1 мы знаем, что у Джона есть отец, Генрих. Вот второй документ (Документ #2):

```
@prefix : <http:> :</http:> henrich <http:> :hasBrother :han
{ ?a :hasFather ?b . ?b :hasBrother ?c . } => { ?a :hasUncle ?c }
</http:>
```

Тут написано, что у Генрихаесть брат Хан. А нижняя строчка это правило, которое говорит, что если у отца некого человека А есть брат, то он дядя для персоны А. Объединяем Документ #1 и #2, соблюдаем правило, и вуаля: компьютер знает, что у Джона есть дядя Хан!

Ричард Сиганиак высказал основные места применения RDF.

- 1. Вам нужно объединять данные из различных источников, не прибегая к созданию специализированных программ.
- 2. Вам нужно дать другим доступ к вашим данным.
- 3. Вам нужно децентрализовать ваши данные так, чтобы ими всеми не «владел» кто-то один.
- 4. Вам нужно сделать что-то особенное с большими объёмами данных вводить, извлекать, просматривать, анализировать, выполнять поиск, и т.д. Вы хотите создать (либо использовать готовый) универсальный инструмент, который бы позволял вам всё это делать, основываясь на модели данных RDF (имеющей то преимущество, что она не привязана к закрытым технологиям хранения и представления данных в отличие от диалектов СУБД).

В общем, это что-то вроде распределенного объединения иерархической и сетевой БД.

Имена в RDF бывают двух типов: литералы (просто текст) и URI. URI может быть ссылкой на веб сайт (http://futuri.us) но не обязательно. В целом URI это ссылка на какой-либо объект (не обязательно он имеет свое представление), к примеру urn:isbn:5-3180-0093-2. Это ссылка на книгу "Samba. Руководство системного администратора. Для профессионалов" Эда Бруксбэнка. У этой книги вряд ли есть страничка, но четко и ясно, что этот URI указывает на книгу, и ясно на какую. URI уникально. Потому позволяет привязывать RDF к какому-то единственному объекту.

```
@prefix : <www.example.org/>
<futuri.us/> a :Website
<futuri.us/> dc:title "Футуриус"
<futuri.us/> dc:contributor "Коробко Ник"
<futuri.us/> dc:creator "Коробко Ник"
<futuri.us/> dc:language ru-RU
<futuri.us/> dc:rights by-nc-sa
```

Краткое описание моего сайта при помощи пространства имен «Дублинское Ядро». Но о нем мы поговорим позднее.

В итоге:

- RDF используется для описания знаний
- описывается тройками значений «подлежащие-сказуемое-объект»
- имеет несколько форм записи: XML и N3
- Использует для значение литералы или Уникальные Идентификаторы (URI)

Полезно:

- В википедии
- W3C/RDF
- Нотация 3
- Английская википедия

Все это с моего блога

Теги: RDF, Semantic Web, Будущее тут, Семантическая сеть

Хабы: Семантика

Комментарии 71



_____ ргохірег 24 декабря 2007 в 05:18 💢

↑ +1 J

Спасибо! Хорошо, что с примерами, а то всё, что попадалось раньше только в каких-то пространственных описаниях (специально, впрочем, информацию не искал).

Какие темы планируете брать дальше? И как насчет описания того, как это можно использовать уже сейчас, не для галочки, а для дела?

Ответить



🌉 Africa 24 декабря 2007 в 12:50 🗰 📮 🔓 🕒



"Какие темы планируете брать дальше? И как насчет описания того, как это можно использовать уже сейчас, не для галочки, а для дела?"

Вот с этим у семнатистов как раз большая проблема. Мне это любопытно и уже несколько раз брался понять, а в чём польза, где практическое применение. И ничиго. Фонтаны перспектив, удивительных возможностей, примитивные примеры... скучно.

Вообще мне это очень сильно напоминает благополучно почившие в бозе экспертные системы. От них также много ждали, а получили ноль. Может где-то и используются, но это не то, что ожидалось.

Моё мнение, семантик веб - это пузырь, поддерживаемый слепыми интузиастами и не слепыми и расчётливыми исследователями, собирающих пыльцу, пока цветёт и никто не понял, что это обман.

Если ошибаюсь - хоть кто-нибудь меня разубедите.

Ответить



shiko_1st 24 декабря 2007 в 13:36 # 🗍 🖒 🙆





Эксперные системы, конечно же, используются. Например, для освоения грантов, выделяемых на их создание:-)

Ответить

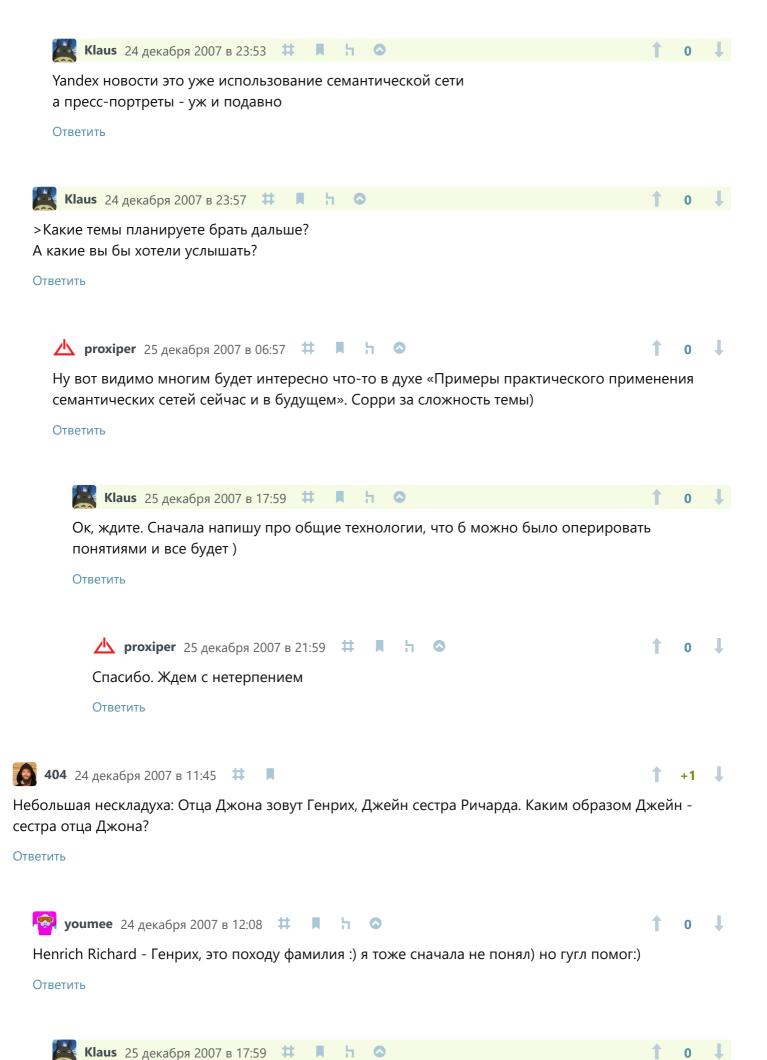


🖍 morhveus 24 декабря 2007 в 15:52 👯 🗍 🖒 🔕





Спасибо, что напомнили про экспертные системы, а то у меня после прочтения возникло лёгкое дежавю.

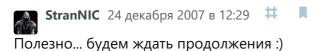


Ответить **уоитее** 25 декабря 2007 в 18:01 🗰 📘 🔓 💿 0 :)) бывает.. но статья интересная Ответить **Klaus** 25 декабря 2007 в 20:34 📫 Спасибо: буду стараться Ответить **↑** -2 ↓ соо19 24 декабря 2007 в 12:10 не для чайников(не понял зачем это все надо Ответить **хаbk** 24 декабря 2007 в 12:11 # **↑** +3 ↓ > описывается тройками значений "подлежащие-прилагательное-сказуемое" На английском (вики/en) — [..] in the form of subject-predicate-object expressions. [..] Subject — подлежащее, predicate — сказуемое, object — дополнение.. где прилагательное?:) Исправьте, пожалуйста. А в целом статья нормальная, спасибо:) Ответить Mastyf 24 декабря 2007 в 12:18 🗰 0 А насколько такой формат быстрее (N3) парсить чем XML? Ответить korchasa 24 декабря 2007 в 14:27 🗰 📘 🔓 🖎 В любом случае парсер узкозаточенный на N3 будет быстрее чем XML'евский, из-за меньшего

словаря, и более простых правил разделения данных (пробелы и концы узлов).

Ответить

нет, тут действительно моя ошибка)



Ответить



кіvsіак 24 декабря 2007 в 12:41 🗰

0

Реальные примеры использования можно у мозиллы посмотреть. Геко как я понял имеет встроенную поддержку RDF. XUL приложения в качестве источников данных во всю его используют.

Ответить



На сколько я смотрел, РДФ в Мозилле хранит журнал.

Ответить



glader 24 декабря 2007 в 12:43 👯 📙

0

+2

Я делал сайт с хранением информации в таких тройках. Очень быстро начинаются проблемы с необходимостью хранения 4-х позиций и т.д. Например "Студия" "выпустила_фильм" "фильм" "в_таком_то_году".

Ответить



Это можно заменить на 2 тройки:

"Студия" "выпустилаФильм" "Фильм"

"Фильм" "вышел" "вТакомТоГоду"

Кстати, с точки зрения организации хранения данных это будет, пожалуй, и более логично.

Ответить



neithere 24 декабря 2007 в 14:13 #

H H G

0

Вот другие примеры:

- В таком-то году Джон был другом Тома.
- Если пойдет дождь И отключат электричество, Мэри скажет "ой".

Иногда выражение имеет смысл только в определенном контексте (1..n дополнительных условий).

Здесь не обойтись триплетом, к сожалению.



Ну да, в таких сложных фразах, конечно, тройкой не обойдёшься. Однако RDF и не предполагает возможности полного семантического разбора естественно-языковых конструкций. Речь идёт об удобном методе хранения данных, когда вместо пары имязначение используется тройка имя-отношение-значение.

Ответить



Дело не в языковых конструкциях, а в том, что они описывают. К сожалению, я не вижу, как на триплеты разложить такую простую фразу как "Джон видел Мэри в 2004 году". Возможно, есть смысл триплет заменить на квадруплет: объект, предикат, субъект и контекст. В этом случае контекст будет массивом других выражений. Возможно, мы как-то так и мыслим...



```
дружба1 — это — ПроцессДружбы
дружба1 — началась — в таком то году
дружба1 — закончилась — в таком то году
дружба1 — в ней учавствовал — джон
дружба1 — в ней учавствовал — том
УсловноеДействие1 — это — УсловноеДействие
УсловноеДействие1 — кто — Мэри
УсловноеДействие1 — что сделает — Скажет1
Скажет 1 — это — Действие Говорения
Скажет1 — что — «ой»
УсловноеДействие1 — перечень условий — Условия1
Условия1 — это — список
Условия1 — первый элемент — Условие1
Условия 1 — остальные элементы — Условия 2
Условия2 — это — список
Условия2 — первый элемент — Условие2
Условия2 — остальные элементы — нету
Условие1 — это — УсловиеЕслиБудетДействие
Условие1 — действие — ПойдётДождь1
ПойдётДождь1 — это — ИзменениеПогоды
ПойдётДождь1 — это когда погода станет — Дождь
ПойдётДождь1 — где — ОколоМэри1
ОколоМэри1 — это — ОтносительноеПоложение1
ОтносительноеПоложение1 — это — ОтносительноеПоложение
ОтносительноеПоложение1 — как — ЗаОкномПомещения
```

```
ОтключатЭлектричество1 — ресурс — Электричество
            Отключат Электричество 1 — доступность кому — Мэри
            Ответить
       ги7701 21 февраля 2017 в 12:58 # 📕 🗎 💿
       Да, да, нужно что-то вроде RDF+контексты. Так мы учтем не только тройки подлежащие-
       сказуемое-дополнение, но и обстоятельства. Желательно туда же добавить и определения.
       Ответить
      0
                                                                                         1
  Тут как-раз и проявляется разница между RDF и relational DB. C RDF должно быть так:
  Студия выпустила_фильм фильмID
  фильмІD называется Ирония_Судьбы_3
  фильмІD выпущен_в 2008
  Ответить
   Daymz 24 декабря 2007 в 12:47 💢
                                                                                      0
Ещё есть хороший формат Turtle (гуглить "RDF Turtle").
В нём ещё сильнее проявляется близость к LISP, благодаря его формату для rdf:first/rdf:rest (они же
лисповские CAR и CDR).
Ответить
атуот 24 декабря 2007 в 13:01 👯
Разъяснитие мне - чайнику для начала что такое WEB 3.0
Ответить
      Maksym 24 декабря 2007 в 13:41 #
  Наверное, то до чего расширится Web2.0:)
  Ответить
    antono 25 декабря 2007 в 04:18
                                                                                    +1
```

ОтносительноеПоложение1 — ОтносительноКого — Мэри

Отключат Электричество 1 — это — Изменение Доступности Ресурсов

Условие2 — это — УсловиеЕслиБудетДействие Условие2 — действие — ОтключатЭлектричество1

Это когда роботы общаются с роботами :)

Ответить



flint 25 декабря 2007 в 18:39



0

Просто громкое модное слово. Для привлечения внимания в основном

Ответить



Klaus 26 декабря 2007 в 18:36 🗰







Не согласен. Кому как. Я использую, чтобы описать новый виток развития сети.

Ответить



Maksym 24 декабря 2007 в 13:14



Спасибо. Очень интересная и полезная информация. Ради такого и читаю Хабр.

Но чтобы это стандарт начал повсеместно использоваться — его должны в первую очередь начать обрабатывать роботы поисковых систем. Есть ли инфа о поддержке с их стороны?

Ответить

НЛО прилетело и опубликовало эту надпись здесь









Вопрос в том, когда она начнет его кушать сама... И соответственно появится смысл вставлять в свой контен описание инфы в формате RDF.

Ответить

НЛО прилетело и опубликовало эту надпись здесь



🦍 **Maksym** 24 декабря 2007 в 17:47 👯





Согласен. Тут рекурсивная зависимость. Но стандарт появился — "значит это кому-нибудь нужно"(с) и процесс наверняка скоро начнет набирать обороты.

Ответить



Klaus 26 декабря 2007 в 18:38 👯

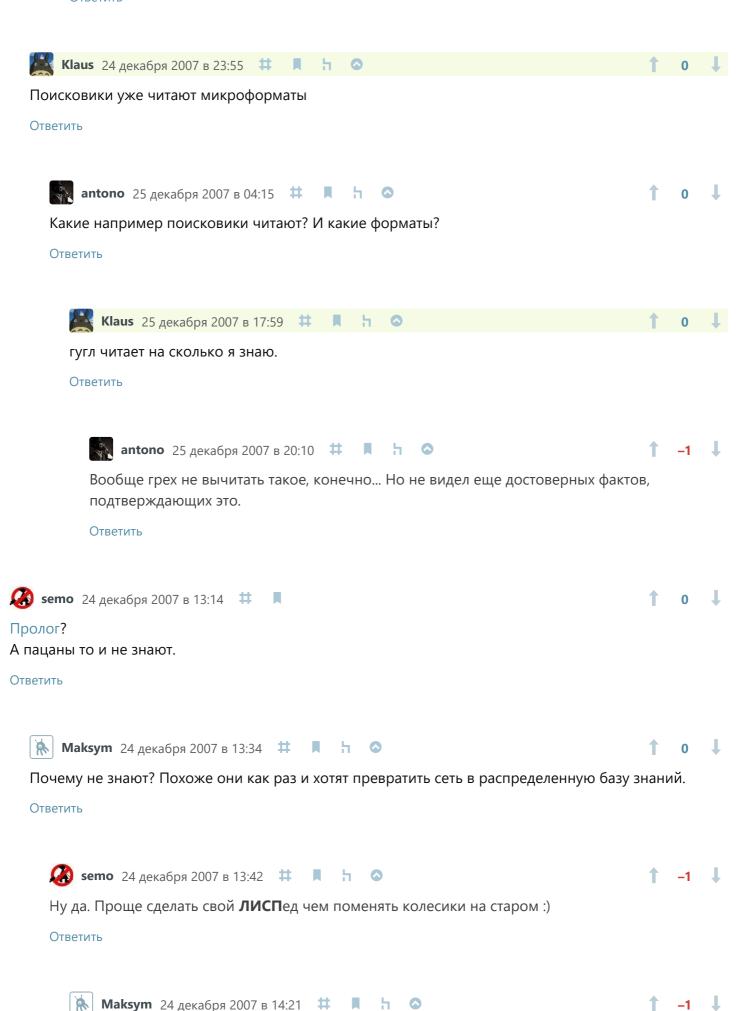






А вы просто включайте.

Это как с Линукосм "не перейду на него, пока не будут писать в основном под него". разработчики: "Не будем писать под линукс, пока все не перейдут"



:D Классный каламбурег! :)

Помимо языков программирования, заточенных именно под логику предикатов, есть еще масса реализаций Rules Engines (знания описываются на RuleML и т.п.) для более распространенных технологий(типа Java или Net), которые тоже неплохо ложатся на обработку RDF.

Ответить







Prolog, во-первых, оперирует предложениями более сложной структуры, а, во вторых, включает в себя не только схему описания отношений, но и схему их анализа. Не стоит путать декларативный язык программирования (Prolog) с языком описания фактов (RDF).

Ответить

НЛО прилетело и опубликовало эту надпись здесь



kronos 24 декабря 2007 в 13:24





Мне кажецо что то подобное широко применяется в экспертных системах. Почитайте про них, наверно понимание придет:)

Ответить



🥰 kronos 24 декабря 2007 в 13:23 👯





Немного похоже на то как турбо пролог сохранял факты.

Ответить



bag 24 декабря 2007 в 13:59 #



В самом начале опечаточка "совй цикл".

И спасибо за разъяснения, но все-равно материал "не для новичков", имхо.

Ответить



Klaus 26 декабря 2007 в 18:39 🗰







Ну нужно знать что такое XML и все такое. А так, поняла даже моя подруга - журналистка.

Ответить



edelweard 24 декабря 2007 в 14:06 #





Граждане из комментариев, это материал для новичков в RDF. То есть для тех, кто не слышал или мало слышал о RDF, но хочет узнать его поближе. Эта статья не для тех, кто ничего не хочет знать вообще. Вы не поняли или вам показалось не интересно, — можете идти дальше, значит, статья не для вас.

Автору большое спасибо. Не критикуя постов другого рода, всё же очень радуюсь появлению вот таких вот конкретных статей по существу.

Ответить



🔅 **ESQUELETO** 24 декабря 2007 в 15:16 🗰 📮



Автору спасибо, но надеюсь, это не выживет... Это личное мнение. Я терпеть не могу функциональное программирование и базы знаний. Хотя, декларативные БЗ не такие уж мозголомальные штуки.

Ответить



Gospodin 24 декабря 2007 в 15:53 #



Сие описанное местами очень напоминает язык Пролог. Пролог сам по себе штука мощная и продуманная, НО(большое но) в нем все завязанно на рекурсии, т.е. сам интерпритатор рекурсивно проверяет факты и строит новые отношения так и стоящие программы без рекурсии организовать не получится средствами Пролога, вот и имеем рекурсивную рекурсию, которая работает не очень быстро, но если это вкурить то нужная логика реализуется весьма просто.

Ответить



🍳 rg_software 24 декабря 2007 в 18:43 💢 📙



Не хочу устраивать холивар, но кто-нибудь вообще верит, что семантик веб будет? Я вот не верю в упор! Да никто не будет аккуратно прописывать метаданные. Или будет прописывать не так, как нужно, а так, как выгодно. Естественно, можно написать про свой сайт, что он лучший в мире, что автомобили "жигули" — вершина мысли, что на сайте можно скачать видео, варез, музыку, фотографии — а на самом деле там лишь домашний хомяк васи пупкина... Т.е. в теории я понимаю, что так можно разметить страничку, и если в пределах одного (ну, одного десятка) сайтов эта концепция будет реализована — это ОК, в это верю. Но весь веб на семантику не переедет, и гугль по метаданным искать не будет. А если будет — то напрасно :)

Ответить



Helg 24 декабря 2007 в 19:59 #





Согласен. Повторится история с keywords и description :)

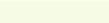
Ответить



Klaus 24 декабря 2007 в 23:56 👯

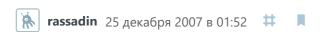








А это не будет так использоваться, это для социального графа будет использоваться.



+2 ↓

Для всех неверующих — очнитесь — оно уже рядом.

Вы, наверное, слышали и даже пользуетесь RSS и FOAF, а это и есть первые словари RDF, т.е это информация, уже описанная по правилам семантичкого веба.

Ответить



0

Еще один интересный пример - http://ontoworld.org/wiki/Berlin

Ответить

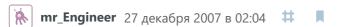


煤 sintez 25 декабря 2007 в 17:02 👯

+2

Спасибо, полезная статья, ждем продолжения! Я как раз сейчас диплом начинаю писать, суть которого в создании semantic-web проекта.

Ответить



Отличная статья. Натолкнула на мысль о необходимости создания фреймворка для работы с RDF потоками. Что то вроде Lisp on rails или Prolog on rails по тому же принципу что Ruby on rails и Grails. В итоге к уже проработанным и оттестированным Java компонентам, отвечающим за работу сервисов, будут добавлены логические компоненты с формальным описанием решения для задачи.



> Самое сильное отличие RDF от XML заключается в том...

Вы сейчас сравниваете круглое с красным. RDF ортогонален XML'у, у которого своей собственной семантики практически нет, один только голый синтаксис. И это не "синтаксис похож", а тот самый синтаксис и есть, плюс семантика RDF.

Ответить



0

Существует принципиальная разница между двумя понятиями: семантическая сеть (Semantic Network) и семантическая паутина (Semantic Web).

В данной статье идет как раз подмена понятий, описана семантическя паутина, а выдается она как семантическая сеть.

Ответить



0

А как сказать что у Класса XXX есть свойсво Y?

XXX rdfs:subClassOf ParentClass XXX ??? Y

что вместо вопросиков ставить?

Ответить



нужно написать Y rdfs:domain XXX

(CM. www.w3.org/TR/2014/REC-rdf-schema-20140225/#ch_property)